



PATENT ABSTRACTS OF JAPAN

(11) Publication number: **2002244547 A**

(43) Date of publication of application: 30.08.02

(51) Int. Cl.

G09B 19/04
G06F 17/60
G09B 5/04
G09B 19/00
G10L 13/00
G10L 21/06
G10L 21/04
G10L 19/00

(21) Application number: 2001041875

(22) Date of filing: 19.02.01

(71) Applicant: **NIPPON HOSO KYOKAI <NHK>**

(72) Inventor: **TSUGI TORU**
HIGUCHI JUNICHI
ABE TEI
NAKAMURA AKIRA
IKEZAWA TATSU
YOKOO TADAAKI

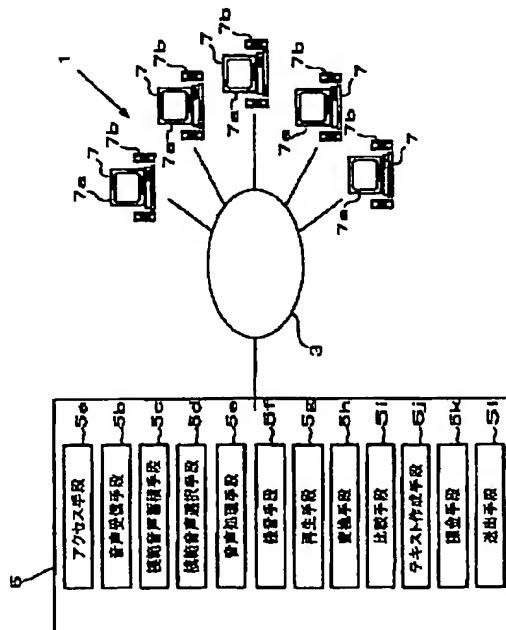
(54) **COMPUTER PROGRAM FOR UTTERANCE
 LEARNING SYSTEM AND SERVER DEVICE
 COLLABORATING WITH THE PROGRAM**

(57) Abstract:

PROBLEM TO BE SOLVED: To provide a computer program utilizing a network which can instantly change the subject of a model speech to enable continuous learning and to provide a server device collaborating with the program.

SOLUTION: The computer program for utterance learning utilizing a network consists of a network 3, a server device 5 connected with the network 3 to store model speeches uttered by a plurality of announcers to be used as a teaching material for mastering correct utterance, and terminals 7 connected with the server device 5 via the network 3 for learning utterance utilizing the model speeches.

COPYRIGHT: (C)2002,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2002-244547
(P2002-244547A)

(43) 公開日 平成14年8月30日 (2002.8.30)

(51) Int.Cl. ⁷	識別記号	F I	テーマコード(参考)
G 0 9 B 19/04		G 0 9 B 19/04	2 C 0 2 8
G 0 6 F 17/60	1 2 8	G 0 6 F 17/60	1 2 8
	3 3 2		3 3 2
	5 0 2		5 0 2
G 0 9 B 5/04		G 0 9 B 5/04	

審査請求 有 請求項の数 8 O L (全 10 頁) 最終頁に続く

(21) 出願番号 特願2001-41875(P2001-41875)

(22) 出願日 平成13年2月19日 (2001.2.19)

(71) 出願人 000004352

日本放送協会

東京都渋谷区神南2丁目2番1号

(72) 発明者 都木 徹

東京都世田谷区砧一丁目10番11号 日本放送協会 放送技術研究所内

(72) 発明者 樋口 淳一

東京都渋谷区神南2丁目2番1号 日本放送協会 放送センター内

(74) 代理人 100064414

弁理士 磯野 道造

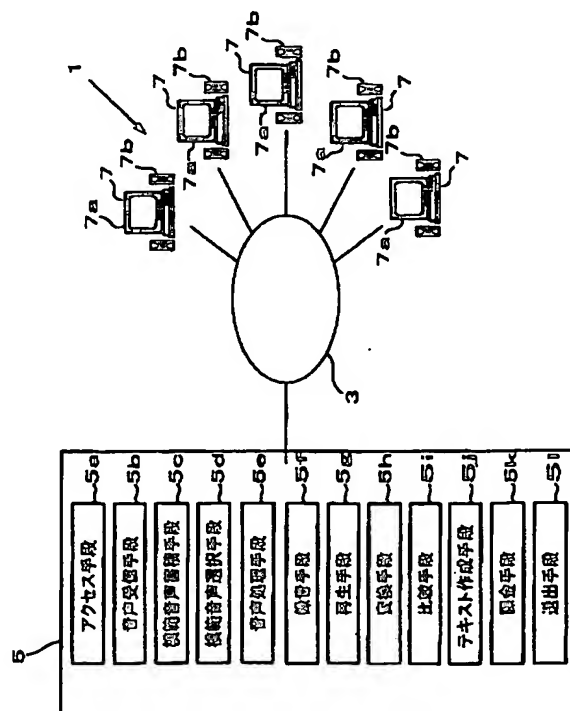
最終頁に続く

(54) 【発明の名称】 発話学習システムのコンピュータプログラムおよびこのプログラムと協働するサーバ装置

(57) 【要約】

【課題】 即時に模範音声の題材を変更することができ、継続して学習を行えるネットワークを利用した発話学習システムのコンピュータプログラムおよびこのプログラムに協働するサーバ装置を提供する。

【解決手段】 ネットワーク3と、ネットワーク3に接続され、正確な発話を修得するための教材用に、複数のアナウンサーが発話した模範音声を蓄積しているサーバ装置5と、ネットワーク3を介してサーバ装置5に接続され、模範音声を利用して発話学習に利用する端末7とを備えたネットワークを利用した発話学習システムのコンピュータプログラムである。



【特許請求の範囲】

【請求項1】 サーバ装置およびこのサーバ装置にネットワークを介して接続される端末を、

正確な発話を修得するための学習教材用であるアナウンサーが発話した模範音声をサーバ装置に蓄積する模範音声蓄積手段、

前記ネットワークを介して前記端末から前記サーバ装置にアクセスするアクセス手段、

前記端末を利用する端末利用者が前記模範音声であるアナウンサーの音声から所望の音声を検索して、選択する模範音声選択手段、

選択された模範音声を、ネットワークを介して端末に送信する送信手段、

送信された模範音声を端末に表示出力する表示出力手段、として機能させるための発話学習システムのコンピュータプログラム。

【請求項2】 端末を利用する端末利用者の発話した音声を入力する音声入力手段、

入力された音声を録音する録音手段、

録音された音声を再生する再生手段、

入力され、録音された端末利用者の音声を処理する音声処理手段、

音声処理された端末利用者が発話した音声を表示出力する表示出力手段、を付加した請求項1に記載の発話学習システムのコンピュータプログラム。

【請求項3】 前記音声処理手段によって処理された音声の抑揚、話速、声質の少なくとも一つを模範音声及び／又は任意の音声の抑揚、話速、声質に変換する変換手段、を付加した請求項2に記載の発話学習システムのコンピュータプログラム。

【請求項4】 前記表示出力手段は、模範音声及び／又は任意の音声を表示すると共に、前記端末利用者が発話した音声を表示し、この表示出力手段によって表示された結果を、前記端末利用者が加工して、この加工した結果に基づいた抑揚、話速、声質に模範音声及び／又は任意の音声を変換する変換手段を付加した請求項2または請求項3に記載の発話学習システムのコンピュータプログラム。

【請求項5】 端末利用者の音声と、前記模範音声選択手段に蓄積されている模範音声によって作成されたサンプル文とに基づきテキストを作成するテキスト作成手段、を付加した請求項2乃至請求項4に記載の発話学習システムのコンピュータプログラム。

【請求項6】 前記端末利用者が発話した音声と前記模範音声とを比較し、評価する比較評価手段、を付加した請求項2乃至請求項5に記載の発話学習システムのコンピュータプログラム。

【請求項7】 前記模範音声蓄積手段に蓄積されている各模範音声を利用することに対する料金が予め設定され、料金を算出する課金手段を付加した請求項1乃至請

求項6に記載の発話学習システムのコンピュータプログラム。

【請求項8】 端末利用者が利用する端末からネットワークを介してアクセスされるサーバ装置であって、正確な発話を修得するための学習教材用であるアナウンサーが発話した模範音声をサーバ装置に蓄積する模範音声蓄積手段と、

前記端末利用者が前記模範音声であるアナウンサーの音声から所望の音声を検索して、選択する模範音声選択手段と、

選択された模範音声を、ネットワークを介して端末に送信する送信手段と、

を備えたことを特徴とするサーバ装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、学習システムを提供するものであり、語学（母国語、外国語の発話）を修得するための模範音声を聴覚で聴取すると共に、視覚で把握し、繰り返し練習できるネットワークを利用した発話学習システムに関するもので、この発話学習システムのコンピュータプログラムおよびこのプログラムと協働するサーバ装置に関する。

【0002】

【従来の技術】一般に、ネイティブスピーカ等が発話した模範音声をテープ、CD等の記憶媒体に記憶させ、この記憶媒体が、当該記憶媒体を購入した利用者の語学修得における発話学習に供されている。

【0003】また、語学教室等において、この教室に通う学習者は、当該語学教室の教師が発声した模範音声を聴取し、これを真似て発声し、語学教室の教師は、学習者の発声した音声を講評する、といった聴覚中心の学習が行われている。

【0004】

【発明が解決しようとする課題】しかしながら、従来の記憶媒体に記憶された模範音声は、この模範音声を発声した者がネイティブスピーカであっても、発話の正式な訓練を受けているとは限らず、記憶媒体を購入した利用者が、この記憶媒体を発話学習の学習教材にした場合に、発声した者（ネイティブスピーカ）の独特の癖等も身に付いてしまうという問題がある。また、従来の記憶媒体に記憶された模範音声は、リアルタイムの話題に関係なく、利用者の興味を引く内容でない場合、何度も聴取しているうちに、利用者は模範音声に飽きてしまう恐れがあり、利用者の語学修得の意欲を減衰させてしまうといった問題がある。また、模範音声は、発声者が不明な（有名でない）場合が多く、利用者は、この音声では親しみがわかず、語学修得の継続に支障をきたすといった問題がある。

【0005】また、語学教室での学習は聴覚中心であるので、学習者は、自分の発声した音声と教師の発声した

音声との差が具体的に把握できない、という問題がある。しかも、学習者の中には、一回の発話に対し、一回の講評を得ることに嫌悪感を覚えることも少なくなく、繰り返し発話の練習を行いたい学習者には不向きな側面があるという問題がある。

【0006】本発明の目的は上述した従来の技術が有する課題を解消し、特に発話の訓練を受けたアナウンサーが発話した模範音声进行学习教材として有効活用し、即時に模範音声の題材を変更することができ、継続して学習を行えるネットワークを利用した発話学習システムのコンピュータプログラムおよびこのプログラムと協働するサーバ装置を提供することにある。

【0007】

【課題を解決するための手段】請求項1記載の発明は、サーバ装置およびこのサーバ装置にネットワークを介して接続される端末を、正確な発話を修得するための学習教材用であるアナウンサーが発話した模範音声サーバ装置に蓄積する模範音声蓄積手段、前記ネットワークを介して前記端末から前記サーバ装置にアクセスするアクセス手段、前記端末を利用する端末利用者が前記模範音声であるアナウンサーの音声から所望の音声を検索して、選択する模範音声選択手段、選択された模範音声を、ネットワークを介して端末に送信する送信手段、送信された模範音声を端末に表示出力する表示出力手段、として機能させるための発話学習システムのコンピュータプログラムに関するものである。

【0008】かかる構成によれば、端末を利用する端末利用者は、ネットワークを介してサーバ装置にアクセスして、サーバ装置に蓄積されている模範音声であるアナウンサーの音声を選択でき、選択されたアナウンサーの音声は端末に表示出力される。尚、アナウンサーが発話する音声において、発話する人数は1人であっても複数人数であってもよく、好ましくは複数人数のアナウンサーが発話する方がよい。

【0009】模範音声とは、特定の単語を発声したものだけではなく、一定の長さの文章を読んだ（発話した）ものも含まれる。

【0010】請求項2記載の発明は、請求項1に記載の発話学習システムのコンピュータプログラムにおいて、端末を利用する端末利用者の発話した音声を入力する音声入力手段、入力された音声を録音する録音手段、録音された音声を再生する再生手段、入力され、録音された端末利用者の音声を処理する音声処理手段、音声処理された端末利用者が発話した音声を表示出力する表示出力手段、を付加したものである。

【0011】かかる構成によれば、端末を利用する端末利用者は、録音手段および再生手段によって、自ら発話した音声録音され再生され、音声処理手段によって音声処理がされ、さらに、表示出力手段によって表示出力される。これらの手段により、自ら発話した音声は繰り

返し聴取され、この自ら発話した音声を表示出力手段を利用して表示することで、視覚で把握される。

【0012】請求項3記載の発明は、請求項2に記載の発話学習コンピュータプログラムにおいて、前記音声処理手段によって処理された音声の抑揚、話速、声質の少なくとも一つを模範音声及び／又は任意の音声の抑揚、話速、声質に変換する変換手段、を付加したものである。

【0013】かかる構成によれば、端末利用者の音声の抑揚、話速、声質の少なくとも一つを模範音声及び／又は任意の音声の抑揚、話速、声質に変換されるので、この変換された音声は、端末利用者の発話したい目標の発話学習用のものとして利用される。尚、任意の音声とは、複数の端末利用者が発話した音声、或いは、予め、アナウンサー以外の人（発話の訓練を受けていない人等）が発話した音声を模範音声蓄積手段に蓄積しておいたもの或いは端末の録音手段を用いて録音しておいたものである。

【0014】請求項4記載の発明は、請求項2または請求項3に記載の発話学習システムのコンピュータプログラムにおいて、前記表示出力手段は、模範音声及び／又は任意の音声を表示すると共に、前記端末利用者が発話した音声を表示し、この前記表示出力手段によって表示された結果を、前記端末利用者が加工して、この加工した結果に基づいた抑揚、話速、声質に模範音声及び／又は任意の音声を変換する変換手段を付加したものである。

【0015】かかる構成によれば、表示出力手段によって、模範音声及び／又は任意の音声或いは、端末利用者が発話した音声が表示され、変換手段によって、表示出力手段による表示結果を、端末利用者が加工した結果に基づいた抑揚、話速、声質に模範音声及び／又は任意の音声を変換される。

【0016】請求項5記載の発明は、請求項2乃至請求項4に記載の発話学習システムのコンピュータプログラムにおいて、端末利用者の音声と、前記模範音声選択手段に蓄積されている模範音声によって作成されたサンプル文とに基づきテキストを作成するテキスト作成手段、を付加したものである。

【0017】かかる構成によれば、端末利用者が発話した音声に基づき、サーバ装置に保持されているサンプル文からテキストが作成されるので、端末利用者の好みに応じた発話学習用の教材が得られる。

【0018】請求項6記載の発明は、請求項2乃至請求項5に記載の発話学習システムのコンピュータプログラムにおいて、前記端末利用者が発話した音声と前記模範音声とを比較し、評価する比較評価手段、を付加したものである。

【0019】かかる構成によれば、端末利用者が発話した音声模範音声に近似しているかを比較評価されるの

で、端末利用者の発話した音声、どの程度模範音声に近づいたかが、端末利用者側で把握できる。

【0020】請求項7記載の発明は、請求項1乃至請求項6に記載の発話学習システムのコンピュータプログラムにおいて、前記模範音声蓄積手段に蓄積されている各模範音声を利用することに対する料金が予め設定され、料金を算出する課金手段を付加したものである。

【0021】かかる構成によれば、端末利用者は各模範音声の予め設定された利用料金の支払いをする。例えば、模範音声の種類や取得量、或いは利用した時間に応じて料金の支払いがされる。

【0022】請求項8記載の発明は、端末利用者が利用する端末からネットワークを介してアクセスされるサーバ装置であって、正確な発話を修得するための学習教材用である複数のアナウンサーが発話した模範音声をサーバ装置に蓄積する模範音声蓄積手段と、前記端末の利用者が前記模範音声を選択する模範音声選択手段と、選択された模範音声を、ネットワークを介して端末に送信する送信手段と、を備えたことを特徴とするサーバ装置に関するものである。

【0023】かかる構成によれば、端末を利用する利用者は、サーバ装置の模範音声蓄積手段に蓄積されているアナウンサーが発話した模範音声、模範音声選択手段によって選択され、端末利用者の要求に応じて送出される送出手段によって、ネットワークに送出される。この模範音声をリアルタイムの話題に随時更新してゆけば、常に新しい模範音声配信される。

【0024】尚、端末は表示出力手段を備えてもよい。この表示出力手段は、サーバ装置に蓄積されている音声を、例えば、時間と1モーラの幅とを横軸に、周波数変化を図示するといった形で表示するものである。また、端末はマイク、A/D変換器等からなる音声入力手段を備えてもよい。そして、端末又はサーバ装置は、録音手段、再生手段、音声処理手段、比較評価手段を備えてもよい。この音声処理手段は、利用者の発声する音声を処理すると共に、実際に模範音声と利用者の発声する音声とを比較する、音声信号処理ソフトと、そのソフトを作動可能にする動作パラメータとからなるソフトウェアである。

【0025】また、サーバ装置は、端末利用者から入力された音声の抑揚、話速、声質のうち少なくとも1つを模範音声及び／又は任意の音声の抑揚、話速、声質に変換する変換手段や模範音声蓄積手段に模範音声で生成されたサンプル文を保持し、端末利用者から入力された音声に基づき、サンプル文が選択され、テキストが作成されるテキスト作成手段やサーバ装置内に蓄積された模範音声には予め各音声毎に料金が設定され、前記端末が前記サーバ装置から取得した模範音声の量に応じて課金する課金手段を備えてもよい。

【0026】さらに、サーバ装置には、端末利用者が発話した音声を認識し、認識した音声に基づきサーバ装置が操作される音声認識手段、音声合成手段を備えてもよい。

【0027】音声処理手段を備えた場合には、端末利用者から入力された音声の基本周波数を抽出したりするなどの処理がされる。

【0028】変換手段を備えた場合には、端末利用者から入力された音声の抑揚、話速、声質のうち少なくとも1つが模範音声の抑揚、話速、声質に変換され、端末利用者は、この変換された音声を端末で利用できる。

【0029】比較評価手段を備えた場合には、端末利用者から入力された音声と、模範音声或いは変換手段によって変換された音声とを比較評価され、端末利用者は自分の発声した音声と模範音声との具体的な差を認識できる。

【0030】ここで、音声処理手段、変換手段、比較評価手段とは、利用者の発声する音声を処理するソフトおよび実際に模範音声と利用者の発声する音声とを比較するソフトウェアであって、このソフトウェアがサーバ装置側にある場合、サーバ装置には音声受信手段が含まれる。

【0031】テキスト作成手段を備えた場合には、端末利用者から入力された音声による指示と、サーバ装置内の模範音声蓄積手段に蓄積されているサンプル文とによって、テキスト（教材、手本）が作成される。ここでサンプル文とは、例えば、国語辞典等の辞書に記載されている例文や、ニュース番組、新聞等で用いられる文章などの様々な文章である。

【0032】課金手段を備えた場合には、模範音声には予め各音声毎に料金が設定され、端末がサーバ装置から取得した模範音声の量に応じて課金され、端末利用者の要求に応じて模範音声配信され、端末利用者は適切な使用料金が支払える。

【0033】ここで、この明細書中、発声、発話によって生じるものは音声であって、その生じる音声の長さによって使い分けることとする。実質的に、発声、発話の語彙に差違はないものとする。また、声質とは、母音や子音の音色のことをいう。

【0034】

【発明の実施の形態】以下、本発明の一実施形態を図面に基づいて詳細に説明する。

（発話学習システム）図1は発話学習システムの概略を示した説明図である。図1に示すように発話学習システム1は、ネットワーク3とサーバ装置5と端末7とを備えている。

【0035】ネットワーク3は、この実施の形態では、汎用プロトコルであるTCP/IPが使用されるインターネットであって、このネットワーク3には、特定の場所、例えば、放送局内に設置されたサーバ装置5が接続

されると共に、不特定多数の場所、例えば、語学学習者の自宅（家庭）、学校（大学、専門学校も含む）、研究機関、図書館等に設置された複数の端末7が接続されている。

【0036】また、端末7は、図1に示したコンピュータの態様に限らず、ゲーム機等のコントローラでもよく、端末7の設置場所は日本国内に限らず、海外の不特定多数の場所であってもよい。つまり、この発話学習システムは、端末利用者の年齢を問わず、端末利用者がコンピュータ等を操作ができなくとも容易に利用でき、さらには、発話学習システムは国内に住む人だけではなく、海外に住む日本人或いは日本語以外の言語を母国語にする人々も対象に作成されている。

【0037】サーバ装置5は、一般に流通しているクライアントサーバシステムにおけるサーバ装置であって、図示を省略した主制御装置、記憶装置、入出力装置、通信制御装置等がデータ制御線によって接続される構成を有している。この実施の形態では、サーバ装置5は、アクセス手段5aと、音声受信手段5bと、模範音声蓄積手段5cと、模範音声選択手段5dと、音声処理手段5eと、録音手段5fと、再生手段5gと、変換手段5hと、比較評価手段5iと、テキスト作成手段5jと、課金手段5kと、送出手段5lとを備えて構成されている。

【0038】アクセス手段5aは、サーバ装置5がネットワーク3上に開示しているホームページを指しており、端末利用者は、端末7からこのホームページを介して発話学習システム1を利用する。

【0039】音声受信手段5bは、上記した入出力装置（図示せず）に相当し、ネットワーク3を介して端末7から入力された音声を受信するものである。

【0040】模範音声蓄積手段5cは、上記した記憶装置（図示せず）に相当し、模範音声（アナウンサーの音声等）と模範音声によって作成されたサンプル文とを蓄積するものである。また、この模範音声蓄積手段5cは、さらに、イントネーション、アクセント辞典を備えている。

【0041】模範音声選択手段5dは、端末からの要求に基づいて、模範音声蓄積手段5cに蓄積されているアナウンサー等の音声から要求された音声を検索して、選択するものである。つまり、この模範音声選択手段5cでは、検索ツールが備えられており、所望のアナウンサーの音声における、特定の音声（例えば、基礎会話、長文発話、著名な作家の小説等）を検索することができる。

【0042】音声処理手段5eは、主制御装置（図示せず）に起動するプログラムであって、受信した音声を分析して、音声の基本周波数を抽出したり、有声音と無声音との区間を識別したりするものである。また、この音声処理手段は、分析結果を端末利用者からの要求に基づ

いて、表示する際に表示調整する機能を有している。

【0043】録音手段5fと再生手段5gは、音声受信手段5bが受信した音声を一時的に録音（保存）し、端末利用者からの要求に基づいて、録音した音声を再生するものである。

【0044】変換手段5hは、主制御装置（図示せず）に起動するプログラムであって、端末利用者を選択された模範音声（アナウンサーの声等）または別の端末利用者が発話した任意の音声を参照して、音声処理された音声の抑揚（音の高低）、話速、声質を、模範音声の抑揚、話速、声質に変換するものである。そして、端末利用者が当該発話学習システム1によって発話学習した結果、この変換された音声に到達できなかった際（うまく真似ることができなかった際）には、この変換された音声、つまり、変換された音声は端末利用者の発話した音声の特徴（抑揚、話速、声質）を変換したものを、そのまま活用することにより、例えば、冠婚葬祭のスピーチを行う場合に、スピーチを行う者は、当該変換された音声を携帯の録音装置に録音しておき、変換された音声を再生すると共に、いわゆる口パクでスピーチを行う、ということができる。

【0045】また、変換手段5hによって、任意の音声を参照して、端末利用者が発話し音声処理された音声の抑揚（音の高低）、話速、声質を、任意の音声の抑揚、話速、声質に変換した場合、例えば、ベテランのアナウンサーの音声の抑揚（音の高低）、話速、声質を新米のアナウンサーの音声の抑揚、話速、声質に変換すると、新米のアナウンサーにとっては、ベテランアナウンサーの音声で、ベテランアナウンサーの音声の特長と自分の音声の特長とを比較することができ、自分の発話した音声の問題点を容易に把握できる。

【0046】比較評価手段5iは、主制御装置（図示せず）に起動するプログラムであって、端末利用者から入力され音声処理された音声と、模範音声或いは変換された音声との抑揚、話速、声質とを比較し、比較した結果、端末利用者と模範音声とどれだけ差があるのかを端末7に伝達するものである。この実施の形態では、後記するように、端末7が表示出力手段を備えているので、端末7の端末利用者は自分の発声した音声と模範音声との差が具体的に視覚で確認できる。

【0047】テキスト作成手段5jは、主制御装置（図示せず）に起動するプログラムであって、端末利用者の発声した音声に基づいて、模範音声蓄積手段5cに蓄積されているサンプル文を参照して端末利用者が所望するテキスト文章を作成するものである。例えば、端末利用者の発声した音声に基づいて、模範音声蓄積手段5cに蓄積されている数種類のサンプル文が端末7に表示され、そのサンプル文には予め序文、本文結びの文ごとにブロック化されており、各サンプル文には番号が付与されている。端末利用者は、各サンプル文の番号を入力す

るだけで所望するテキストが自動的に作成される。また、このテキスト作成手段5jは、例えば、端末7に備えられたワープロソフト等を利用して作成された文章データ（テキストファイル）から、模範音声蓄積手段5cに蓄積されている模範音声を使用してテキストを作成できるプログラムである。

【0048】課金手段5kは、主制御装置（図示せず）に起動するプログラムであって、模範音声蓄積手段5cに模範音声を蓄積する際に、この蓄積される各音声に料金を設定すると共に、端末7の端末利用者が端末7に取得した模範音声の量に応じて料金を算出するものである。

【0049】送出手段5lは、上記した入出力装置（図示せず）に相当し、前記した模範音声、再生された音声、変換された音声、比較した結果、テキスト文章、利用料金のデータ等をネットワーク3に送出するものである。

【0050】端末7は、一般に流通しているクライアントサーバシステムにおけるクライアントにあたり、すなわち、端末7はサーバ装置5に対し、様々な処理を要求する信号を送信するものであり、表示手段7a、出力手段7bと、図示を省略した音声入力手段とを備えて構成されている。ここで、表示手段7aと出力手段7bとを合わせて請求項に記載の表示出力手段とする。

【0051】ここで、様々な処理を要求する信号とは、模範音声蓄積手段5cに蓄積されている模範音声から端末利用者が特定の模範音声を選択する模範音声選択信号や、端末利用者が発声した音声の変換を要求する変換要求信号や、端末利用者が発声した音声と模範音声との比較を要求する比較要求信号や、模範音声蓄積手段5cに蓄積されているサンプル文によってテキスト文章の作成を要求するテキスト作成要求信号等のことである。

【0052】表示手段7aは、ディスプレイ等で構成され、サーバ装置5から送信された上記した模範音声、変換された音声、比較した結果、テキスト文章等を表示するものである。音声入力手段（図示せず）は、マイク等で構成され、端末利用者の発声した音声を入力するためのものである。出力手段（音声出力手段）7bは、スピーカ等で構成され、サーバ装置5から送信された模範音声等を出力するためのものである。

【0053】（発話学習システムの動作）次に、発話学習システム1の動作を、図2を参照して説明する。この発話学習システム1では、サーバ装置5側に音声処理手段5eが備えられている。この音声処理手段5eは、端末7側に設けられてもよい。

【0054】サーバ装置5は、ネットワーク3に対し、ホームページ等のアクセス情報を開示しており、このアクセス情報に端末7がアクセスすると（S1）、サーバ装置5は、端末7から模範音声蓄積手段5cに蓄積されている模範音声を選択される（模範音声選択信号を受信す

る）まで待機する（S2）。

【0055】サーバ装置5は、端末7から模範音声選択信号を受信すると、模範音声蓄積手段5cに蓄積されている模範音声の中から選択された模範音声を主制御装置（図示せず）で処理可能な状態にする（S3）。

【0056】次に、サーバ装置5は、端末7から端末利用者の発声した音声が入力まで待機し（S4）、入力されたら音声受信手段で端末利用者の発声した音声を受信する（S5）。受信された端末利用者の発声した音声は音声処理手段5eで基本周波数が抽出される等の処理がされる（S6）。

【0057】そして、サーバ装置5は、端末7からの変換要求信号を受信したかどうかを判断する（S7）。受信した場合には、端末利用者が発声した音声を選択された模範音声の抑揚または話速或いは声質に変換する（S8）。変換要求信号を受信しない場合には、このS8は省略される。

【0058】次に、サーバ装置5は、端末7からの比較要求信号を受信したかどうかを判断する（S9）。受信した場合には、端末利用者が発声した音声と、模範音声或いは変換された音声との抑揚、話速、声質等を比較する（S10）。比較要求信号を受信しない場合には、このS11は省略される。

【0059】さらに、サーバ装置5は、端末7からのテキスト作成要求信号を受信したかどうかを判断する（S11）。受信した場合には、端末利用者が発声した音声と、模範音声蓄積手段5cによって蓄積されているサンプル文とに基づいて、端末利用者が要求するテキスト文章を作成する（S12）。テキスト作成要求信号を受信しない場合には、このS12は省略される。

【0060】そして、サーバ装置5は、課金手段5kによって、端末7が取得した模範音声の量に応じて料金を算出する（S13）。

【0061】この後、送出手段5lによって、模範音声や変換された音声や比較結果やテキスト文章および発話学習システム1を利用した料金の情報がネットワーク3を介して端末7に送られる（S14）。

【0062】（動作例）次に、発話学習コンピュータプログラムの動作例を説明する。図3は日本語の発話学習コンピュータプログラムによって端末7のディスプレイ上に表示された模範音声「ひろいうみをみる」を示したものである。この図3は、横軸に時間を取り、縦軸に声の高さを示したものである。

【0063】この図3において、模範音声は、実線で示され、端末利用者が発声した音声は、破線で示されている。ディスプレイの右下方の部分には、図中左から順に再生、長短変換、上下変換（周波数変換）、停止、メニューを示すアイコンが表示されている。

【0064】端末利用者は、このアイコンを端末7に付随しているマウス等のデバイス（図示せず）を使用し

て、これらのアイコンをクリックする。

【0065】「再生」のアイコンをクリックすると、「模範音声の再生」、「自分の音声の再生」、「任意の音声の再生」「模範音声に基づいた変換音声の再生」「任意の音声に基づいた変換音声の再生」等のスクロールメニューが表示され、さらに「模範音声の再生」を選択すると再生する速度が表示される。この速度を選択すると、模範音声或いは任意の音声はその速度に従って発声される。

【0066】また、表示されたスクロールメニュー中の「模範音声に基づいた変換音声の再生」をクリックすると、端末利用者が発声した音声の抑揚、話速、声質のいずれかを模範音声の抑揚、話速、声質に変換するかが問われ、例えば、抑揚を選択すると、端末利用者が発声した音声の抑揚が、模範音声の抑揚に変換される。同様に、「任意の音声に基づいた変換音声の再生」をクリックすると、例えば、任意の音声「早口で語尾の上がる音声」であった場合、端末利用者が発声した音声「早口で語尾の上った音声」に変換される。つまり、端末利用者の発声した音声の話速が上がり、語尾の抑揚（音声周波数）が上がった音声に変換される。

【0067】「長短変換」のアイコンをクリックすると、ディスプレイ上のマウスのポインタが、例えば、矢印から十字に変化する。そして、その十字のポインタによりディスプレイ上の特定の1カ所を指定し、クリック後、ドラッグすると、ドラッグした部分の音声を引き延ばされる。

【0068】「上下変換（周波数変換）」のアイコンをクリックすると、同様に、ディスプレイ上のマウスのポインタが、例えば、矢印から十字に変化する。そして、その十字のポインタによりディスプレイ上の特定の1カ所を指定し、クリック後、上下移動させる（ドラッグする）と、ドラッグした部分の音声周波数の高低が変化する。例えば、「はし（橋）」（「は」の方にアクセントがある）という音声の「し」の部分を上記したように操作して、周波数を高くすると「はし（箸）」（「し」の方にアクセントがある）といったように同音異義の音声にすることができる。すなわち、端末利用者の好み等によって、端末7のディスプレイ上で加工した結果に基づいて、指定した音声に変換される。これによれば、端末利用者は、好みの音声を自在に加工することで、楽しみながら発話学習に取り組むことができる。

【0069】「停止」のアイコンをクリックすると、即座に発声（出力）中の音声の停止される。

【0070】「メニュー」のアイコンをクリックすると、「評価」「次の音声」「はじめに戻る」等のスクロールメニューが表示される。表示されたスクロールメニュー中の「評価」をクリックすると、自分の発声した音声と模範音声とがどのように違うかを評価してくれる。例えば、「ひろいうみ」の「い」の部分を『心持ち下げ

て発声するように』というように評価がされる。

【0071】この実施の形態では以下の効果を奏す。端末7を利用する端末利用者は、サーバ装置5に蓄積されている発話の訓練を受けたアナウンサーが発話した模範音声を利用して発話学習を行えるので、正確な発話を修得することができる。また、この模範音声をリアルタイムの話題に随時更新してゆけば、端末利用者は飽きることなく発話学習を継続することが可能になる。

【0072】端末7の端末利用者の好みに応じてアナウンサーの音声を選択できるので、端末7の端末利用者は、発話の訓練を受けたアナウンサーの音声を学習教材にして、独特の癖のない標準的な発話を身につけることができ、発話学習に親しみを持って取り組むことができる。

【0073】端末利用者の発声した音声の抑揚、話速、声質の少なくとも1つが、変換手段5hによって、模範音声の抑揚、話速、声質に変換されるので、端末利用者はこの変換された音声を学習教材として利用することができ、この学習教材は、端末7の端末利用者の発話したい目標の音声になる。このため、端末7の端末利用者は目的意識を持って継続して発話学習を行うことができる。また、端末利用者が当該発話学習システム1によって発話学習した結果、この変換された音声に到達できなかった際には、この変換された音声を、そのまま活用することができる。

【0074】端末7の端末利用者が発声した音声と模範音声に近似しているかを比較評価手段5iによって比較されるので、端末7の端末利用者の発声した音声と、どの程度模範音声に近づいたかが、端末利用者側で把握される。端末7に表示出力手段（図示せず）を備えている場合には、聴覚だけではなく、視覚によって端末利用者が発声した音声と模範音声との差が具体的に把握されるので、端末利用者は目的意識を持って発話学習に取り組むことができる。

【0075】端末利用者が発声した音声に基づき、サーバ装置5の模範音声蓄積手段5cに保持されているサンプル文からテキストが作成されるので、端末利用者の好みに応じた学習教材を得ることができ、この学習教材を利用して発話練習をした音声の評価がされるので発話学習の学習意欲を維持することができる。

【0076】また、端末7の端末利用者が、このサーバ装置5に蓄積されているサンプル文を使用すれば、例えば、端末利用者が結婚式のスピーチ原稿を作成する際の手間を省くことができる。このため、端末利用者は、所望のスピーチ原稿を、好みの模範音声（アナウンサーの音声等）で作成されたものをテキストとして利用することができる。

【0077】端末7の端末利用者は、課金手段によって算出された、模範音声の取得量に応じて料金を支払うので、従来のテープ、CD等の記憶媒体を購入するより

も、端末利用者は必要に応じた模範音声に対し適切な料金を支払うことができる。以上、一実施形態を説明したが、実施形態はこれに限定されるものではない。

【0078】

【発明の効果】請求項1記載の発明によれば、端末を利用する利用者は、サーバ装置に蓄積されている模範音声を利用して発話学習を行えるので、この模範音声をリアルタイムの話題に随時更新してゆけば、利用者は飽きることなく発話学習を継続することが可能になる。

【0079】請求項2記載の発明によれば、模範音声にアナウンサーの音声を用いられ、しかも、端末利用者の好みに応じてアナウンサーの音声を選択できるので、端末利用者は、発話の訓練を受けたアナウンサーの音声を学習教材にして、独特の癖のない標準的な発話を身につけることができ、発話学習に親しみを持って取り組むことができる。

【0080】請求項3記載の発明によれば、端末利用者の音声の抑揚、話速、声質の少なくとも1つを模範音声及び／又は任意の音声の抑揚、話速、声質に変換されるので、端末利用者はこの変換された音声を学習教材として利用ことができ、この学習教材は、端末利用者の発話したい目標の音声になる。このため、端末利用者は目的意識を持って継続して発話学習を行うことができる。

【0081】請求項4記載の発明によれば、表示出力手段によって、模範音声及び／又は任意の音声或いは、端末利用者が発話した音声が表示され、変換手段によって、表示出力手段による表示結果を、端末利用者が加工した結果に基づいた抑揚、話速、声質に模範音声及び／又は任意の音声に変換されるので、端末利用者は、模範音声や任意の音声を、端末利用者の好み等により、自在に変換でき、楽しみながら発話学習に取り組むことができる。

【0082】請求項5記載の発明によれば、端末利用者が発声した音声が発話音声に近似しているかを評価されるので、端末利用者の発声した音声、どの程度模範音声に近づいたかが、端末利用者側で把握される。端末に表示機能を備えている場合には、聴覚だけではなく、視覚によって端末利用者が発声した音声と模範音声との差が具体的に把握されるので、端末利用者は目的意識を持って発話学習に取り組むことができる。

【0083】請求項6記載の発明によれば、端末利用者が発声した音声に基づき、サーバ装置に保持されているサンプル例文からテキストが作成されるので、端末利用者の好みに応じた学習教材を得ることができ、この学習教材を利用して発話練習をした音声の評価されるので発話学習の学習意欲を維持することができる。

【0084】また、端末利用者が、このサーバ装置に蓄

積されているサンプル例文を使用すれば、例えば、端末利用者が結婚式等のスピーチ原稿を作成する際の手間を省くことができる。このため、端末利用者は、所望のスピーチ原稿を、好みの模範音声（アナウンサーの音声）で作成されたものをテキストとして利用することができる。

【0085】請求項7記載の発明によれば、端末利用者は、各模範音声の予め設定された利用料金の支払いをする。例えば、模範音声の種類や取得量、或いは利用した時間に応じて料金の支払いがされるので、従来のテープ、CD等の記憶媒体を購入するよりも、端末利用者は必要に応じた模範音声に対し適切な料金を支払うことができる。

【0086】請求項8記載の発明によれば、端末を利用する利用者は、サーバ装置の模範音声蓄積手段に蓄積されている模範音声、端末利用者の要求に応じて送出される送出手段によって、ネットワークに送出され、端末で利用して発話学習を行うことができる。この模範音声をリアルタイムの話題に随時更新してゆけば、利用者は飽きることなく発話学習を継続することが可能になる。

【図面の簡単な説明】

【図1】本発明によるネットワークを利用した発話学習システムの説明図である。

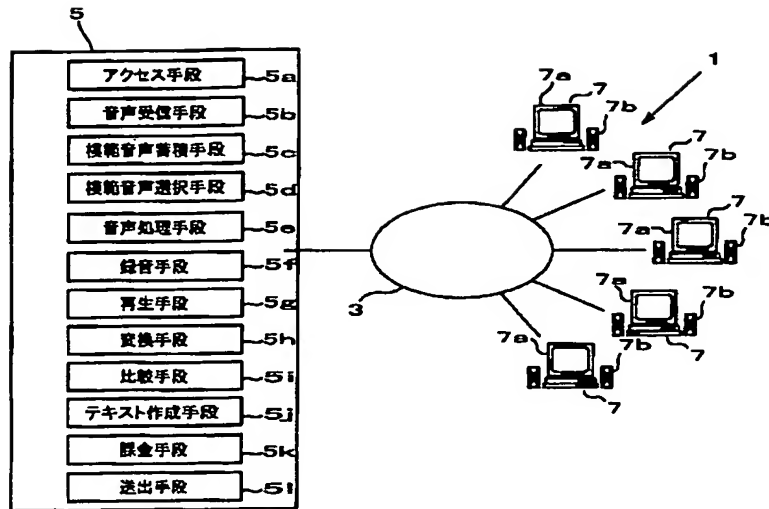
【図2】ネットワークを利用した発話学習システムのフローチャートである。

【図3】発話学習システムの動作例を説明する説明図である。

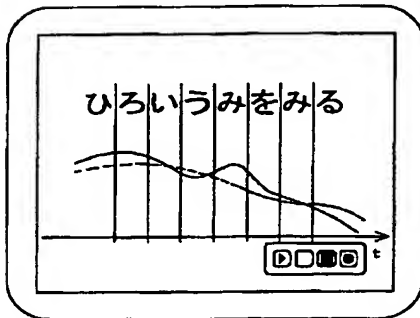
【符号の説明】

- 1 発話学習システム
- 3 ネットワーク
- 5 サーバ装置
- 5 a アクセス手段
- 5 b 音声受信手段
- 5 c 模範音声蓄積手段
- 5 d 模範音声選択手段
- 5 e 音声処理手段
- 5 f 録音手段
- 5 g 再生手段
- 5 h 変換手段
- 5 i 比較手段
- 5 j テキスト作成手段
- 5 k 課金手段
- 5 l 送出手段
- 7 端末
- 7 a 表示手段
- 7 b 出力手段

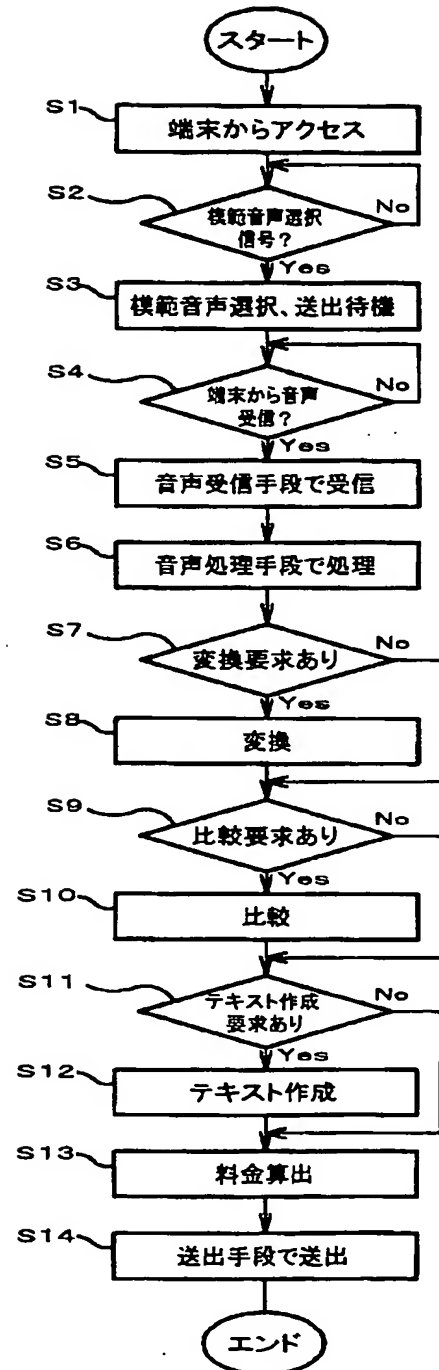
【図1】



【図3】



【図2】



フロントページの続き

(51) Int. Cl. 7

G 0 9 B 19/00

G 1 0 L 13/00

識別記号

F I

G 0 9 B 19/00

G 1 0 L 3/00

メモコード (参考)

H

S

21/06

3/02

A

21/04

7/10

19/00

9/18

J

(72) 発明者 阿部 悌
秋田県秋田市山王 1-1-2 日本放送協
会 秋田放送局内

(72) 発明者 中村 章
東京都渋谷区神南二丁目 2 番 1 号 日本放
送協会 放送センター内

(72) 発明者 池沢 龍
東京都渋谷区神南二丁目 2 番 1 号 日本放
送協会 放送センター内

(72) 発明者 横尾 忠晃
東京都渋谷区神南二丁目 2 番 1 号 日本放
送協会 放送センター内

F ターム (参考) 2C028 AA01 BA01 BA03 BB04 BB06
BC02 BD02 CA13 CB03 CB04
5D045 AB13 BA01 BA02 DB01